Topological Data Analysis Made Easy with the Topology ToolKit, What is New?

To cite this version:
Martin Falk, Christoph Garth, Charles Gueunet, Pierre Guillou, Attila Gyulassy, et al.. Topological Data Analysis Made Easy with the Topology ToolKit, What is New?. IEEE VIS Tutorials, Oct 2020, Salt Lake City, United States. hal-02950538

HAL Id: hal-02950538
https://hal.science/hal-02950538
Submitted on 28 Sep 2020
Topological Data Analysis Made Easy with the Topology ToolKit,
What is New?

Martin Falk
Linköping University

Christopher Garth
TU Kaiserslautern

Charles Gueunet
Kitware

Pierre Guillou
Sorbonne Universite

Attila Gulyassy
University of Utah

Lutz Hofmann
Heidelberg University

Christopher Kappe
TU Kaiserslautern

Julien Tierny
CNRS, Sorbonne Universite

Joshua A. Levine
University of Arizona

Jules Vidal
Sorbonne Universite

Joshua A. Levine
University of Arizona

Jonas Lukasczyk
Arizona State University

Figure 1: TTK is a software platform for topological data analysis in scientific visualization. It is both easily accessible to end users (ParaView plugins (a), VTK-based generic GUIs (b) or command-line programs (c)) and flexible for developers (Python (d), VTK/C++ (e) or dependence-free C++ (f) bindings). TTK provides an efficient and unified approach to topological data representation and simplification, which enables in this example a discrete Morse-Smale complex (a) to comply to the level of simplification dictated by a piecewise linear persistence diagram (bottom-right linked view, a). Code snippets are provided (d-f) to reproduce this pipeline.

1 LEVEL OF THE TUTORIAL
This tutorial is targeted at a Beginner audience.

2 POTENTIAL SCHEDULE CONFLICTS
If possible, we would like to avoid any scheduling overlap with IEEE LDAV 2020 which takes place on October 25. Coauthor Tierny is a co-chair of LDAV this year, and many of the speakers, as well as the potential attendees, are members of both communities.

3 POTENTIAL FOR AN ONLINE EVENT
In the case IEEE VIS 2020 had to be an online event, we would run the tutorial virtually. For this, we would be happy to adapt our organization to the recommendations provided by the IEEE VIS organizers. For instance, we could replicate the organization successfully experimented at EuroVis 2020 (mandatory pre-registration, talk pre-recording, online chatting with Discord, Zoom for live presentation capture, YouTube for content broadcasting). Also, one of the tutorial members would act as master of ceremony and would be in charge of running the Discord channel and to transfer all questions/remarks to the tutorial presenters, who would also actively contribute to the online chat.

4 ABSTRACT
This tutorial presents topological methods for the analysis and visualization of scientific data from a user’s perspective, with the Topology ToolKit (TTK), an open-source library for topological data analysis. Topological methods have gained considerably in popularity and maturity over the last twenty years and success stories of established methods have been documented in a wide range of applications (combustion, chemistry, astrophysics, material sciences, etc.) with both acquired and simulated data, in both post-hoc and in-situ contexts. This tutorial aims to fill a gap by providing a beginner’s introduction to topological methods for practitioners, researchers, students, and lecturers. In particular, instead of focusing on theoretical aspects and algorithmic details, this tutorial focuses on how topological methods can be useful in practice for concrete data analysis tasks such as segmentation, feature extraction or tracking. The tutorial describes in detail how to achieve these tasks with TTK. In comparison to the last two iterations of this tutorial [11, 13], this iteration emphasizes the features of TTK which now appear to be the most popular, as well as the latest additions to the library. First, we provide a general introduction to topological methods and their application in data analysis, and a brief overview of TTK’s main entry point for end users, namely ParaView, will be presented. Second, we will proceed to a hands-on session demoting the main features of TTK as well as its most recent additions. Third, we will present advanced usages of TTK, including the usage of TTK with Python, the development of a new module for TTK as well as the integration of TTK into a pre-existing system. Presenters of this tutorial include experts in topological methods, core authors of TTK as well as active users, coming from academia and industry. A large part of the tutorial will be dedicated to hands-on exercises and a rich material package will be provided to the participants. This tutorial mostly targets students, practitioners and researchers who are not necessarily experts in topological methods but who are interested in using them in their daily tasks. We also target researchers already familiar to topological methods and who are interested in using or contributing to TTK. We kindly ask potential attendees to optionally pre-register at the following address, in order for us to reach out to them ahead of the tutorial with information updates (for instance, last minute updates, instructions for the download of the tutorial material package, etc.): https://forms.gle/CvY3oWZB9hWSQIb9

Tutorial web page (including all material, TTK pre-installs in virtual machines, code, data, demos, video tutorials, slides, etc): https://topology-tool-kit.github.io/ieeeVisTutorial.html

5 TUTORIAL ORGANIZATION
Motivations Topological analysis techniques [24,33,39] have shown to be practical solutions in various contexts: isosurface extraction [7,34,45], feature tracking [36], volume rendering [49], data simplification [43] and compression [37], similarity estimation [12,44],
geometry processing [40, 47]. They enable the concise and complete capture of the structure of the input data into high-level topological abstractions such as contour trees [6, 16, 17], Reeb graphs [19, 31, 32, 42], or Morse-Smale complexes [9, 21, 22]. Successful applications in a variety of fields of science have been documented (combustion [4, 20, 28], fluid dynamics [5, 8, 26], material sciences [14, 23, 29], chemistry [2, 15, 30], and astrophysics [35, 38]), which further demonstrates the importance of these techniques.

While reference textbooks have been published [10], topological methods have not yet been widely adopted as a standard data analysis tool. We believe one of the reasons for this is the lack of open-source software that implement these algorithms in a generic, user-friendly, and easy-to-use way. The Topology ToolKit (TTK) [3, 41] has been released (BSD license) to fill this gap and 15 institutions have contributed to its development so far. Since its release in 2017, TTK’s website collected 210k page-views, from 23k visitors. This indicates that a user base exists and that further efforts towards the explanation of TTK’s usage would be beneficial to the community.

Target audience This tutorial mostly targets beginners, students, practitioners, and researchers who are not experts in topological methods. It also targets researchers already familiar to topological methods and who are interested in using or contributing to TTK.

Tutorial goals The goals of this tutorial are to present the key tools in topological data analysis (the Persistence diagram, the Reeb graph and its variants, the Morse-Smale complex, etc.) and how they can be used in practice for precise data analysis tasks, including data segmentation and feature extraction. All examples will be illustrated with TTK. We expect participants to become capable of using TTK with ParaView independently, after attending the tutorial.

Hands-on material A large part of the tutorial will be dedicated to hands-on exercises with TTK and ParaView [1]. We will provide a rich material package including TTK pre-installs in virtual machines, (to be used by attendees during the tutorial), data, demos, etc. Most of this material is already available on TTK’s website [11, 13]. Our idea is that participants with a laptop will be able to follow along regardless of their native OS. Attendees who attend just to listen and learn will receive sufficient material to try our examples at home.

Optional pre-registration In 2018, we observed that many attendees would benefit from having time before the tutorial to install materials and have their systems set up with the hands-on material. If our proposal is accepted, we will use the following on-line form (https://forms.gle/CvrY3oWZB9hWSQJb9) to notify potential audiences of where and how to download the material and build a mailing list to help form an informal community for the tutorial.

Proposal strengths In contrast to previous tutorials on topological methods [48], we believe this proposal to have a unique concrete and applicable appeal, by its focus on the usage of topological methods rather than on their foundations. Thus, we expect it to attract a larger audience than the specific subset of IEEE VIS attendees typically found in traditional topology sessions. Moreover, in comparison to the previous editions of this tutorial, our hands-on session will put a special emphasis on TTK’s features which have shown to be the most popular up to now, as well as the latest additions to the library.

We believe that the list of presenters is also a strength of this proposal. First, it includes topology experts as well as core developers and users of TTK. More importantly, it includes researchers with a variety of experience profiles (Ph.D. students, post-docs, professors) and backgrounds (industry, academia), which will ease interactions with a potentially heterogeneous audience. Moreover, the particularly large number of presenters has two merits. First, it imposes a mini-symposium structure, where speakers will give presentation lasting between 10 and 20 minutes, which will result in a lively rhythm in the overall tutorial. Second, this large number of presenters will be instrumental during the hands-on exercises, as there will be enough presenters such that one presenter can assist a small group of attendees (typically 3 to 4). Finally, we believe the detailed program of the tutorial (see Sect. 5) achieves a balance between concepts, usage descriptions and application examples.

Detailed content The tutorial is divided into three main parts (each part being subdivided into modules), for a target duration of approximately 3 hours. These three groups of modules can be organized differently to fit any standard structure for breaks to match the tutorial schedule of IEEE VIS. After the tutorial concludes, we will make all content available from our website.

Followup Survey Since this would be the third iteration of this tutorial, and we hope to continue to run it in future years, we also plan to collect survey data from the participants to help provide feedback to the organizers and shape the content and structure of future tutorials, that might be hosted at additional venues.

A. Preliminaries (60 minutes)
A1. General introduction (5 minutes, by Julien Tierny)
A2. Introduction to topological methods for data analysis (30 minutes, Attila Gyulassy) This talk will present the core tools in topological data analysis (the Persistence diagram [10], the Reeb graph and its variants [6, 17–19, 32, 42], the Morse-Smale complex [9, 21, 22]). In particular, it will detail how these tools can be used for data segmentation and feature extraction.
A3. Quick introduction to ParaView’s user interface (25 minutes, by Charles Gueunet) This talk will provide a brief description of ParaView’s main interface [1], in order to support its usage for beginners in the subsequent hands-on session. This will cover the usage of filters, pipeline design, view manipulation, and Python exports.

B. Hands-on exercises (60 minutes)
B1. General usage of TTK (10 minutes, by Julien Tierny) This talk will briefly describe TTK’s usage philosophy. It will briefly present how TTK can be used from ParaView, Python, VTK/C++ or C++.
B2. Running TTK with Docker (10 minutes, by Christoph Garth)
B3. Segmenting medical data with merge trees (10 minutes, by Charles Gueunet) This hands-on TTK/ParaView exercise will be a step-by-step tutorial showing how to extract individual bones in a medical CT scan interactively with merge trees.
B4. Extracting filament structures with the Morse-Smale complex (10 minutes, Pierre Guillou) This hands-on TTK/ParaView exercise will show step-by-step how to extract filament structures with the Morse-Smale complex on chemistry data.

C. Advanced usage (60 minutes)
C1. Using TTK with Python (10 minutes, Latz Hofmann) This talk will present how to deploy (with its Anaconda package) and use TTK in a Python environment with simple examples.
C2. Developing a new module for TTK (10 minutes, Jonas Lakaszyk) This hands-on will present how to create in 2 minutes a new, running module in TTK. This presentation will also cover the main steps of a module creation.
C3. TTK Integration into Inviwo (15 minutes, by Martin Falk) Inviwo [25] is a rapid prototyping framework for data visualization. TTK has recently been integrated into Inviwo to extend it with topological analysis. In this talk, we will provide a brief introduction on the concepts utilized in Inviwo like its network editor and the associated data flow paradigm. We then detail our approach of integrating TTK into the data flow within Inviwo, which involves transforming data structures from Inviwo to TTK and back. The seamless integration is demonstrated with a number of examples.
C4. TTK as a teaching platform (10 minutes, by Joshua Levine) This talk will provide feedback about our experience in using TTK in our topological data analysis classes.
Jules Vidal – jules.vidal@sorbonne-universite.fr – is a Ph.D. student at Sorbonne Université. He received the engineering degree in 2018 from ENSTA ParisTech. His notable contributions to TTK include the efficient and progressive approximation of distances, barcenters and clusterings of persistence diagrams [27, 46].

**ACKNOWLEDGMENTS**

This work is partially supported by the European Commission grant ERC-2019-COG TROTF (ref. 863446). This work is also supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under Award Number DE-SC-0019109.

**REFERENCES**


